GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 6, 2002, 22:41:46 ; Search time 2691.5 Seconds
(without alignments)
16489.608 Million cell updates/sec

Title: US-10-025-514-15
Perfect score: 1525
Sequence: 1 tctagaccatggaagaccct.....ccagtcaaggcctagtcgac 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	629.4	41.3	1308	6	AR111412	AR111412 Sequence
2	628.4	41.2	1185	6	AR111411	AR111411 Sequence
3	437	28.7	1345	9	HUMA1ATB	M11465 Human alpha
4	433.8	28.4	1434	6	E00631	E00631 DNA encodin
5	432.2	28.3	1312	6	I02706	I02706 Sequence 1
6	430.6	28.2	1584	9	BC011991	BC011991 Homo sapi
7	429	28.1	1185	6	AR111410	AR111410 Sequence
8	429	28.1	1352	6	AX335339	AX335339 Sequence
9	429	28.1	1352	9	HUMA1ATM	K01396 Human alpha
10	429	28.1	1371	6	AX335338	AX335338 Sequence
11	429	28.1	1371	9	HSATPR1	X01683 Human mRNA
12	429	28.1	1399	9	AK026174	AK026174 Homo sapi
13	429	28.1	1434	6	E00195	E00195 cDNA encodi
14	429	28.1	1434	6	I04196	I04196 Sequence 3
15	429	28.1	1434	6	I04272	I04272 Sequence 1
16	429	28.1	1434	6	I07849	I07849 Sequence 2
17	429	28.1	1435	6	AX019486	AX019486 Sequence
18	429	28.1	2478	17	AF130068	Af130068 Homo sapi
19	427.4	28.0	1337	9	HUMA1ATZ	J02619 Human Z typ
20	427.4	28.0	1378	6	I02398	I02398 Sequence 8
21	427.4	28.0	1431	9	BC015642	BC015642 Homo sapi
22	427.4	28.0	2571	17	AF113676	Af113676 Homo sapi
23	425.8	27.9	1299	6	I07949	I07949 Sequence 2
24	425.8	27.9	1308	9	BABA1ATA	J00321 Papio hamad
25	425.8	27.9	1317	6	I00556	I00556 Sequence 2
26	425.8	27.9	1378	6	I03509	I03509 Sequence 6
27	425.8	27.9	1378	6	I07947	I07947 Sequence 4
28	425.8	27.9	1434	6	I01352	I01352 Sequence 1
29	425.4	27.9	1185	6	A01846	A01846 Human mRNA
30	424.2	27.8	1378	6	I01227	I01227 Sequence 2
31	417	27.3	1191	9	AB004044	AB004044 Cercopith
32	407	26.7	1390	6	AX202089	AX202089 Sequence
33	404.6	26.5	1356	6	I36163	I36163 Sequence 16
34	403	26.4	1356	6	I36161	I36161 Sequence 12
35	402.6	26.4	2013	6	AX472008	AX472008 Sequence
36	399.8	26.2	1356	6	I36164	I36164 Sequence 18
37	390.2	25.6	1351	10	AB000550	AB000550 Spermophi
38	387	25.4	1372	10	AB000552	AB000552 Callosciu
39	385.4	25.3	1351	4	SSANTIELA	X88780 S.scrofa mR
40	378	24.8	1306	10	RATATRA1	M32247 Rat alpha-1
41	377.2	24.7	1352	10	S77822	S77822 alpha-1-ant
42	376.4	24.7	1380	10	RATA1PI	D00675 Rat mRNA fo
43	372.8	24.4	1343	10	AB000546	AB000546 Tamias si
44	372.6	24.4	1242	6	E13268	E13268 Tamias asia
45	370.6	24.3	1380	4	BTA1AT	X63129 B.taurus mR

ALIGNMENTS

RESULT 1
AR111412
LOCUS AR111412 1308 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 6 from patent US 6127145.
ACCESSION AR111412
VERSION AR111412.1 GI:12828260
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1308)
AUTHORS Sutliff,T.D. and Rodriguez,R.L.
TITLE Production of .alpha. .sub.1 -antitrypsin in plants
JOURNAL Patent: US 6127145-A 6 03-OCT-2000;
FEATURES Location/Qualifiers

```
    source          1. .1308
                    /organism="unknown"
BASE COUNT      298 a    442 c    363 g    205 t
ORIGIN

  Query Match           41.3%;  Score 629.4;  DB 6;  Length 1308;
  Best Local Similarity  70.8%;  Pred. No. 4.8e-124;
  Matches  837;  Conservative    0;  Mismatches  346;  Indels    0;  Gaps    0;

Qy     11 GGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCA 70
          ||| ||||| || |||||||||||| || || |||||||||||| || |||||||| |||||
Db    116 GGAGGACCCGCAGGGCGACGCCGCCCAGAAGACCGACACCAGCCACCACGACCAGGACCA 175

Qy     71 TCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACA 130
          ||||| || || || || || ||||| |||||||| || ||     |||| | ||
Db    176 CCCGACGTTCAACAAGATCACCCCGAATTTGGCCGAATTCGCCTTCAGCCTGTACCGCCA 235

Qy    131 ATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGC 190
          | || || ||    || || || ||||| || || || || ||    || ||||| ||
Db    236 GCTCGCGCACCAGTCCAACTCCACCAACATCTTCTTCAGCCCGGTGAGCATCGCCACCGC 295

Qy    191 TTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTT 250
          |||||||||| ||    | ||||| || || || ||||| |||||||| | |||||  |
Db    296 CTTCGCCATGCTGTCCCTGGGTACCAAGGCGGACACCCACGACGAGATCCTCGAAGGGCT 355

Qy    251 AAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTT 310
          ||||| ||  |||| || ||||| || || || |||||||| || || ||| || | |
Db    356 GAACTTCAACCTGACGGAGATCCCGGAGGCGCAGATCCACGAGGGCTTCCAGGAGCTGCT 415

Qy    311 GAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTT 370
          || ||  | || || || || || ||  | ||  | || || || |||||  | || |
Db    416 CAGGACGCTCAACCAGCCGGACTCCCAGCTCCAGCTCACCACCGGCAACGGGCTCTTCCT 475

Qy    371 GTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAG 430
          ||| || ||  | ||  | || || || ||||| || ||||| ||||| || || ||
Db    476 GTCCGAGGGCCTCAAGCTCGTCGATAAGTTCCTGGAGGACGTGAAGAAGCTCTACCACTC 535

Qy    431 TGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTA 490
          ||||| || ||||| || || || || || ||||| || || ||||| || || || ||
Db    536 CGAGGCGTTCACCGTCAACTTCGGGGACACCGAGGAGGCCAAGAAGCAGATCAACGACTA 595

Qy    491 TGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATAC 550
          || ||||| || |||||||| |||||||| ||||| || || ||||| ||  | || ||
Db    596 CGTCGAGAAGGGGACCCAGGGCAAGATCGTGGACCTGGTCAAGGAATTGGACAGGGACAC 655

Qy    551 CGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGT 610
          ||||||||| || || |||||  || |||||||| |||||||| || || ||||||||
Db    656 CGTCTTCGCGCTCGTCAACTACATCTTCTTCAAGGGCAAGTGGGAGCGCCCGTTCGAGGT 715

Qy    611 TAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAAT 670
          || || || || ||||| || || || || || || || || || ||||| || || ||
Db    716 GAAGGACACCGAGGAGGAGGACTTCCACGTCGACCAGGTCACCACCGTCAAGGTCCCGAT 775

Qy    671 GATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTT 730
          |||||| || || || |||||||| || || || ||||| ||  |      |||||  |
Db    776 GATGAAGAGGCTCGGCATGTTCAACATCCAGCACTGCAAGAAGCTCTCCAGCTGGGTGCT 835

Qy    731 ATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAGGTAAGCT 790
          | ||||||||  | || ||||| || || || ||  | || ||||| || |||||
Db    836 CCTCATGAAGTACCTGGGGAACGCCACCGCCATCTTCTTCCTGCCGGACGAGGGCAAGCT 895

Qy    791 TCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGA 850
          || ||  | ||||| ||| |||| || ||||| || || || ||  | ||||||||||
Db    896 CCAGCACCTGGAGAACGAGCTGACGCACGACATCATCACGAAGTTCCTGGAGAACGAGGA 955

Qy    851 TCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAA 910
          | ||  ||||   || ||||| || ||| | || |||||||| || |||||| | ||
Db    956 CAGGCGCTCCGCTAGCCTCCACCTCCCGAAGCTGAGCATCACCGGCACGTACGACCTGAA 1015

Qy    911 ATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGG 970
          ||  | |||||| | || || || || || ||    ||||| || ||  |    ||
Db   1016 GAGCGTGCTGGGCCAGCTGGGCATCACGAAGGTCTTCAGCAACGGCGCGGACCTCTCCGG 1075

Qy    971 TGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTAT 1030
          || || || || || ||  | ||  |    || || || ||||| || ||  | || ||
Db   1076 CGTGACGGAGGAGGCCCCCCTGAAGCTCTCCAAGGCCGTGCACAAGGCGGTGCTCACGAT 1135

Qy   1031 TGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCAT 1090
          || || ||||| || || || ||||| || ||||||||||| || || || |||  |||
Db   1136 CGACGAGAAGGGGACGGAAGCTGCCGGGGCCATGTTCCTGGAGGCCATCCCCATGTCCAT 1195

Qy   1091 TCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAA 1150
          || || || || || || || || || ||||| || ||||||||||||||||||||| ||
Db   1196 CCCGCCCGAGGTCAAGTTCAACAAGCCCTTCGTCTTCCTGATGATCGAGCAGAACACGAA 1255

Qy   1151 AAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAG 1193
          |||||  | || ||||| ||||| |||||||| || ||||||
Db   1256 GAGCCCCCTCTTCATGGGGAAGGTCGTCAACCCCACGCAGAAG 1298


RESULT 2
AR111411
LOCUS       AR111411                1185 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 3 from patent US 6127145.
ACCESSION   AR111411
VERSION     AR111411.1  GI:12828259
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1185)
  AUTHORS   Sutliff,T.D. and Rodriguez,R.L.
  TITLE     Production of .alpha. .sub.1 -antitrypsin in plants
  JOURNAL   Patent: US 6127145-A 3 03-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1185
                     /organism="unknown"
BASE COUNT      276 a    396 c    334 g    179 t
ORIGIN

  Query Match           41.2%;  Score 628.4;  DB 6;  Length 1185;
  Best Local Similarity  70.7%;  Pred. No. 7.9e-124;
  Matches  836;  Conservative    0;  Mismatches  346;  Indels    0;  Gaps    0;

Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
          || ||||| || |||||||||||| || || |||||||||||| || |||||||| |||||
Db      1 GAGGACCCGCAGGGCGACGCCGCCCAGAAGACCGACACCAGCCACCACGACCAGGACCAC 60

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
          ||||| || || || || || || ||||| |||||||| || ||     |||| | ||
Db     61 CCGACGTTCAACAAGATCACCCCGAATTTGGCCGAATTCGCCTTCAGCCTGTACCGCCAG 120

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
          | || || ||    || || || ||||| || || || ||    || ||||| ||
Db    121 CTCGCGCACCAGTCCAACTCCACCAACATCTTCTTCAGCCCGGTGAGCATCGCCACCGCC 180

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
          |||||||||  ||    | ||||| || || || |||||  |||||||  | |||||  |
Db    181 TTCGCCATGCTGTCCCTGGGTACCAAGGCGGACACCCACGACGAGATCCTCGAAGGGCTG 240

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
          ||||| ||  |||| || ||||| || || || || ||||||||  || || ||| || |
Db    241 AACTTCAACCTGACGGAGATCCCGGAGGCGCAGATCCACGAGGGCTTCCAGGAGCTGCTC 300

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
          || ||  | || || || || ||  | ||  | || || || |||||  | ||  ||
Db    301 AGGACGCTCAACCAGCCGGACTCCCAGCTCCAGCTCACCACCGGCAACGGGCTCTTCCTG 360

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
          || || ||  | ||  | || || || ||||| || ||||| ||||| || || ||
Db    361 TCCGAGGGCCTCAAGCTCGTCGATAAGTTCCTGGAGGACGTGAAGAAGCTCTACCACTCC 420

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
          ||||| || ||||| || || || || || ||||| || || ||||| || || || ||
Db    421 GAGGCGTTCACCGTCAACTTCGGGGACACCGAGGAGGCCAAGAAGCAGATCAACGACTAC 480
```

```
Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
          || ||||| || |||||||| |||||||| ||||| || || ||||| ||  | || |||
Db    481 GTCGAGAAGGGGACCCAGGGCAAGATCGTGGACCTGGTCAAGGAATTGGACAGGGACACC 540

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
          |||||||| || || ||||| || || |||||||| |||||||| || || ||||||||
Db    541 GTCTTCGCGCTCGTCAACTACATCTTCTTCAAGGGCAAGTGGGAGCGCCCGTTCGAGGTG 600

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
          || || || || ||||| || || || || || || || || || || ||||| || || |||
Db    601 AAGGACACCGAGGAGGAGGACTTCCACGTCGACCAGGTCACCACCGTCAAGGTCCCGATG 660

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
          ||||| || || || |||||||| || || || ||||| ||  |        |||||  |
Db    661 ATGAAGAGGCTCGGCATGTTCAACATCCAGCACTGCAAGAAGCTCTCCAGCTGGGTGCTC 720

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAGGTAAGCTT 791
          | |||||||||  | || ||||| || || || || ||  | || ||||| || |||||
Db    721 CTCATGAAGTACCTGGGGAACGCCACCGCCATCTTCTTCCTGCCGGACGAGGGCAAGCTC 780

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
          || ||  | ||||| ||| |||| || ||||| || || || ||  | ||||||||||||
Db    781 CAGCACCTGGAGAACGAGCTGACGCACGACATCATCACGAAGTTCCTGGAGAACGAGGAC 840

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
          | ||   ||||   || ||||| || ||| | || |||||||| || |||||| | ||
Db    841 AGGCGCTCCGCTAGCCTCCACCTCCCGAAGCTGAGCATCACCGGCACGTACGACCTGAAG 900

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
            ||  | ||||||  | || || || || || ||    ||||| || ||  |    ||
Db    901 AGCGTGCTGGGCCAGCTGGGCATCACGAAGGTCTTCAGCAACGGCGCGGACCTCTCCGGC 960

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
          || || || || || ||  | ||  |    || || || ||||| || ||  | || ||
Db    961 GTGACGGAGGAGGCCCCCCTGAAGCTCTCCAAGGCCGTGCACAAGGCGGTGCTCACGATC 1020

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || ||||| || || || ||||| || |||||||||| || || || |||  |||
Db    1021 GACGAGAAGGGGACGGAAGCTGCCGGGGCCATGTTCCTGGAGGCCATCCCCATGTCCATC 1080

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || || || ||||| || ||||||||||||||||||||| ||
Db    1081 CCGCCCGAGGTCAAGTTCAACAAGCCCTTCGTCTTCCTGATGATCGAGCAGAACACGAAG 1140

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAG 1193
           |||||  | || ||||| ||||| |||||||| || ||||||
Db    1141 AGCCCCCTCTTCATGGGGAAGGTCGTCAACCCCACGCAGAAG 1182
```

RESULT 3
HUMA1ATB

LOCUS	HUMA1ATB 1345 bp mRNA linear PRI 30-OCT-1994
DEFINITION	Human alpha-1-antitrypsin mRNA, complete cds.
ACCESSION	M11465
VERSION	M11465.1 GI:177826
KEYWORDS	alpha-1-antitrypsin.
SOURCE	Human liver, cDNA to mRNA.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1345)
AUTHORS	Ciliberto,G., Dente,L. and Cortese,R.
TITLE	Cell-specific expression of a transfected human alpha 1-antitrypsin gene
JOURNAL	Cell 41 (2), 531-540 (1985)
MEDLINE	85176977
PUBMED	2985281
REFERENCE	2 (sites)
AUTHORS	Curiel,D., Brantly,M., Curiel,E., Stier,L. and Crystal,R.G.
TITLE	Alpha 1-antitrypsin deficiency caused by the alpha 1-antitrypsin Nullmattawa gene. An insertion mutation rendering the alpha 1-antitrypsin gene incapable of producing alpha 1-antitrypsin
JOURNAL	J. Clin. Invest. 83 (4), 1144-1152 (1989)
MEDLINE	89198065
PUBMED	2539391
COMMENT	[2] sites; alpha-1-antitrypsin deficiency disease mutations. The Granite Falls (a one base pair deletion), Bellingham (a substitution) and Mattawa (a one base pair insertion) mutations all result in premature transcription termination.

FEATURES Location/Qualifiers

source	1. .1345
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/map="14q32.1"
gene	1. .1345
	/gene="PI"
CDS	15. .1268
	/gene="PI"
	/note="alpha-1-antitrypsin"
	/codon_start=1
	/protein_id="AAA51546.1"
	/db_xref="GI:177827"
	/db_xref="GDB:G00-120-289"
	/translation="MPSSVSWGILLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDHP TFNKITPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEG LNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKKL YHSEAFTVNFGDHEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKWE RPFEVKDTEDEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIFF LPDEGKLQHLENELTHDIITKFLENEDRRSASLHLPKLSITGTYDLKSVLGQLGITKV FSNGADLSGVTEEAPLKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKP FVFLMIEQNTKSPLFMGKVVNPTQK"
variation	562. .564
	/gene="PI"
	/note="acg in wt; ag in null-Granite Falls"
variation	732
	/gene="PI"
	/note="a in wt; t in null-Bellingham"
variation	1141
	/gene="PI"
	/note="ta in wt; tta in null-Mattawa"

BASE COUNT 354 a 381 c 317 g 293 t
ORIGIN

Query Match 28.7%; Score 437; DB 9; Length 1345;
Best Local Similarity 60.6%; Pred. No. 5.2e-83;
Matches 716; Conservative 0; Mismatches 465; Indels 0; Gaps 0;

```
Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || || ||    || || || || || ||
Db      84 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 143

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || ||  | || || || || ||     | ||  | ||
Db     144 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 203

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || ||
Db     204 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 263

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |      | || || || || || || || || || || ||  | || ||  |
Db     264 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 323

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || ||  |  |
Db     324 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 383

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || || ||  | ||  |
Db     384 CGTACCCTAAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 443

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     444 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 503

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
```

```
Db    504 GAAGCCTTCACTGTCAACTTCGGGGATCACGAAGAGGCCAAGAAACAGATCAACGATTAC 563

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551

Db    564 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 623

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611

Db    624 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCTTTTGAAGTC 683

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671

Db    684 AAGGACACCGAGGACGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTCCCTATG 743

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731

Db    744 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTACTG 803

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791

Db    804 CTAATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTACCTGATGAGGGGAAACTA 863

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851

Db    864 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 923

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911

Db    924 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 983

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971

Db    984 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1043

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031

Db   1044 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1103

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091

Db   1104 GACGAGAAGGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCAATGTCTATC 1163

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151

Db   1164 CCCCCAGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1223

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192

Db   1224 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1264


RESULT 4
E00631
LOCUS       E00631                  1434 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  DNA encoding human alpha-1 antitrypsin mutain.
ACCESSION   E00631
VERSION     E00631.1  GI:2168910
KEYWORDS    JP 1986012289-A/1.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1434)
  AUTHORS   Maagaretsuto,W.I. and Guren,H.K.
  TITLE     REGION SPECIFIC MUTATION INDUCTION IN ALPHA-1-ANTI-TRYPSIN
  JOURNAL   Patent: JP 1986012289-A 1 20-JAN-1986;
            CHIMO JIENETEITSUKUSU INC
COMMENT     OS   human
            PN   JP 1986012289-A/1
            PD   20-JAN-1986
            PF   14-MAR-1985 JP 1985051553
            PR   14-MAR-1984 US 84    589410, 07-MAR-1985 US 85    709382 PI
            MAAGARETSUTO WAI INZURII, GUREN HITOSHI KAWASAKI PC
            C12N15/00,A61K35/74,A61K37/64,C07H21/04,C07K15/04, PC
            C12P21/02//A61K37/04,
            PC   C12N9/99,(C12P21/02,C12R1:865);
            CC   strandedness: Both;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   *source: tissue_type=liver;
            FH   Key             Location/Qualifiers
            FH
            FT   CDS             47. .1303
            FT                   /product='alpha-1 antitrypsin' FT   variation
            replace(1190. .1193,'GCG')
            FT                   /note='amino acid 358=Ala'
            FT   variation       replace(1190. .1193,'GTG')
            FT                   /note='amino acid 358=Val'
            FT   variation       replace(1190. .1193,'GGG')
            FT                   /note='amino acid 358=Gly'
            FT   variation       replace(1190. .1193,'TTC')
            FT                   /note='amino acid 358=Phe'
            FT   variation       replace(1190. .1193,'AGA')
            FT                   /note='amino acid 358=Arg'
            FT   variation       replace(1190. .1193,'AGG')
            FT                   /note='amino acid 358=Arg'
            FT   variation       replace(1190. .1193,'AAG')
            FT                   /note='amino acid 358=Lys'
            FT   variation       replace(1142. .1144,'AAG')
            FT                   /note='amino acid 342=Lys'.
FEATURES             Location/Qualifiers
     source          1. .1434
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      369 a    426 c    346 g    293 t
ORIGIN

  Query Match           28.4%;  Score 433.8;  DB 6;  Length 1434;
  Best Local Similarity 60.5%;  Pred. No. 2.5e-82;
  Matches  714;  Conservative    0;  Mismatches  467;  Indels    0;  Gaps    0;

Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71

Db    119 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 178

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131

Db    179 CCAACCTTCAACAAGATCACCCCGAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 238

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191

Db    239 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 298

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251

Db    299 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 358

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311

Db    359 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTGCTC 418

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371

Db    419 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTG 478

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431

Db    479 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 538

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491

Db    539 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 598

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
```

```
Db    599 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 658

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611

Db    659 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 718

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671

Db    719 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 778

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731

Db    779 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCGAGCTGGGTGCTG 838

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791

Db    839 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 898

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851

Db    899 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 958

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911

Db    959 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1018

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971

Db   1019 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1078

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031

Db   1079 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1138

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091

Db   1139 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1198

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151

Db   1199 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1258

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192

Db   1259 TCTCCCCTGTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1299
```

```
RESULT 5
I02706
LOCUS       I02706                 1312 bp ss-DNA     linear   PAT 21-MAY-1993
DEFINITION  Sequence 1 from Patent US 4599311.
ACCESSION   I02706
VERSION     I02706.1  GI:268359
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Kawasaki,G.H.
  TITLE     Glycolytic promotersfor regulated protein expression: protease
            inhibitor
  JOURNAL   Patent: US 4599311-A 1 08-JUL-1986;
            1547 - 16th Ave. East; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="unknown"
BASE COUNT      339 a    368 c    324 g    281 t
ORIGIN

  Query Match           28.3%;  Score 432.2;  DB 6;  Length 1312;
  Best Local Similarity   60.4%;  Pred. No. 5.5e-82;
  Matches  713;  Conservative    0;  Mismatches  468;  Indels    0;  Gaps    0;
```

```
Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71

Db    100 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 159

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131

Db    160 CCAACCTTCAACAAGATCACCCCCAACTTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 219

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191

Db    220 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 279

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251

Db    280 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 339

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311

Db    340 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 399

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371

Db    400 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 459

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431

Db    460 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 519

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491

Db    520 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 579

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551

Db    580 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 639

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611

Db    640 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 699

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671

Db    700 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 759

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731

Db    760 ATGAAGCGTTTAGGCATGTTTAACATCCAGCATTGTAAGAAGCTGTCCAGCTGGGTGCTG 819

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791

Db    820 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 879

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851

Db    880 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 939

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911

Db    940 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 999

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971

Db   1000 AGCGTCCTAGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1059

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031

Db   1060 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1119

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091

Db   1120 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1179
```

```
Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
          |  || || || || || || ||||| || || ||  | ||||| || || || || ||
Db   1180 CGCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1239

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
            ||  | || ||||| || || || || || || || ||
Db   1240 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1280
```

RESULT 6
BC011991
LOCUS BC011991 1584 bp mRNA linear PRI 02-AUG-2001
DEFINITION Homo sapiens, Similar to serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1, clone MGC:9222 IMAGE:3859644, mRNA, complete cds.
ACCESSION BC011991
VERSION BC011991.1 GI:15080498
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1584)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalon@bcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W., Muzny,D.M., Gibbs,R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 21 Row: d Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6855600.

FEATURES Location/Qualifiers
source 1. .1584
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:9222 IMAGE:3859644"
/tissue_type="Ovary, adenocarcinoma"
/clone_lib="NIH_MGC_66"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
CDS 233. .1489
/codon_start=1
/product="Similar to serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1"
/protein_id="AAH11991.1"
/db_xref="GI:15080499"
/translation="MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDH
PTFNKITPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILE
GLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKK
LYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKW
ERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIF
FLPDEGKLQHLENELTHDIITKFLENEDRRSASLHLPKLSITGTYDLKSVLGQLGITK
VFSNGADLSGVTEEAPLKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNK
PFVFLMIDQNTKSPLFMGKVVNPTQK"

```
BASE COUNT      410 a    459 c    379 g    336 t
ORIGIN

  Query Match           28.2%;  Score 430.6;  DB 9;  Length 1584;
  Best Local Similarity  60.3%;  Pred. No. 1.2e-81;
  Matches  712;  Conservative    0;  Mismatches  469;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || || ||    || || || || || ||
Db     305 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 364

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || | || || || || || ||     | || | ||
Db     365 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 424

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || ||
Db     425 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 484

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || || || | || || |
Db     485 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 544

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || || |  |
Db     545 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 604

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || || || ||  |
Db     605 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCTTGTTCCTC 664

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || || ||  | || || || || || ||  | || ||
Db     665 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 724

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || ||||||
Db     725 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 784

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || || ||  | || || ||  | ||  | || ||
Db     785 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 844

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || ||||  | || || || ||
Db     845 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 904

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     905 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 964

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || || || || ||  |      |||||  |
Db     965 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 1024

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || ||  | || || || || || ||
Db    1025 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 1084

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || || ||  | || || || ||
Db    1085 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 1144

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           |  |    ||     | ||  | || ||  |    || || || || || ||  | ||
Db    1145 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1204

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  | || ||  | || || || || || ||    || || || || ||  |    ||
Db    1205 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1264

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || || ||  | || ||
```

```
Db   1265 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1324

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
          || || || || || || || || || || ||||| | || || || || || |||   ||
Db   1325 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1384

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
          || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db   1385 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGACCAAAATACCAAG 1444

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
             ||  | || ||||| || || || || || || || ||
Db   1445 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1485


RESULT 7
AR111410
LOCUS       AR111410                1185 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 2 from patent US 6127145.
ACCESSION   AR111410
VERSION     AR111410.1  GI:12828258
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1185)
  AUTHORS   Sutliff,T.D. and Rodriguez,R.L.
  TITLE     Production of .alpha. .sub.1 -antitrypsin in plants
  JOURNAL   Patent: US 6127145-A 2 03-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1185
                     /organism="unknown"
BASE COUNT      328 a    324 c    283 g    250 t
ORIGIN

  Query Match           28.1%;  Score 429;  DB 6;  Length 1185;
  Best Local Similarity   60.2%;  Pred. No. 2.7e-81;
  Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || ||    || || || || || ||
Db       1 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 60

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || ||  | || || || || ||     | ||  | ||
Db      61 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 120

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || ||
Db     121 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 180

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || || ||  | || ||  |
Db     181 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 240

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || ||  |  |
Db     241 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 300

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || ||  | ||  |
Db     301 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 360

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     361 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 420

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || ||||||
Db     421 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 480

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || ||  | || || ||  | ||  | || ||

Db     481 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 540

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || ||||  | || || || || ||
Db     541 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 600

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || ||| |||
Db     601 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 660

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || || || || ||  |       |||||  |
Db     661 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 720

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || ||  | || || || || || ||
Db     721 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 780

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || ||  | || || || ||
Db     781 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 840

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           |  |    ||      | ||  | || ||  |    || || || || || ||  | ||
Db     841 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 900

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  | || ||  | || || || || || ||    || || || ||  |    ||
Db     901 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 960

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db     961 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1020

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || || ||||  | || || || || |||   ||
Db    1021 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1080

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db    1081 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1140

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
             ||  | || ||||| || || || || || || || ||
Db    1141 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1181


RESULT 8
AX335339
LOCUS       AX335339                1352 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 5848 from Patent WO0194629.
ACCESSION   AX335339
VERSION     AX335339.1  GI:18126058
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 5848 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1352
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      349 a    386 c    325 g    292 t
ORIGIN

  Query Match           28.1%;  Score 429;  DB 6;  Length 1352;
```

```
 Best Local Similarity   60.2%;  Pred. No. 2.6e-81;
 Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || || ||    || || || || || ||
Db      92 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 151

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || || || || || || || || || || || ||
Db     152 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 211

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || ||
Db     212 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 271

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || || || |  | || ||  |
Db     272 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 331

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || |  | |
Db     332 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 391

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || || ||  | ||  |
Db     392 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 451

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || ||  | || || || || || ||  | || ||
Db     452 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 511

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || ||||| |
Db     512 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 571

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || || ||  | || || ||  | ||  | || ||
Db     572 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 631

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || ||||  | || || || ||
Db     632 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 691

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     692 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 751

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || || || || || ||  |       |||||  |
Db     752 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 811

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || || ||  | || || || || || ||
Db     812 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 871

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || ||  | || || || ||
Db     872 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 931

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           | |   ||    | ||  | || ||  |    || || || || || ||  | ||
Db     932 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 991

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
           ||  | || ||  | || || || || || ||    || || || ||  |    ||
Db     992 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1051

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db    1052 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1111

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || |||||  | || || || || |||   ||

Db    1112 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1171

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db    1172 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1231

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
              ||  | || ||||| || || || || || || || ||
Db    1232 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1272
```

```
RESULT 9
HUMA1ATM
LOCUS       HUMA1ATM                1352 bp    mRNA    linear   PRI 01-NOV-1994
DEFINITION  Human alpha-1-antitrypsin mRNA, complete cds.
ACCESSION   K01396
VERSION     K01396.1  GI:177828
KEYWORDS    alpha-1-antitrypsin; antiprotease; antitrypsin.
SOURCE      Human liver, cDNA to mRNA, clones pULB1523 [1] and pTG603 [3].
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1231)
  AUTHORS   Bollen,A., Herzog,A., Cravador,A., Herion,P., Chuchana,P., Vander
            Straten,A., Loriau,R., Jacobs,P. and van Elsen,A.
  TITLE     Cloning and expression in Escherichia coli of full-length
            complementary DNA coding for human alpha 1-antitrypsin
  JOURNAL   DNA 2 (4), 255-264 (1983)
  MEDLINE   84107980
   PUBMED   6319097
REFERENCE   2  (bases 1 to 1352)
  AUTHORS   Colau,B., Chuchana,P. and Bollen,A.
  TITLE     Revised sequence of full-length complementary DNA coding for human
            alpha 1-antitrypsin
  JOURNAL   DNA 3 (4), 327-330 (1984)
  MEDLINE   85026667
   PUBMED   6333329
REFERENCE   3  (bases 95 to 286)
  AUTHORS   Courtney,M., Buchwalder,A., Tessier,L.H., Jaye,M., Benavente,A.,
            Balland,A., Kohli,V., Lathe,R., Tolstoshev,P. and Lecocq,J.P.
  TITLE     High-level production of biologically active human alpha
            1-antitrypsin in Escherichia coli
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 81 (3), 669-673 (1984)
  MEDLINE   84144765
   PUBMED   6322161
COMMENT     Alpha-1-antitrypsin is synthesized in the liver and is a major
            constituent of plasma.  It functions as an inhibitor of elastase,
            which degrades lung tissues.  Hereditary deficiencies are fairly
            frequent, causing early lung degeneration and sometimes severe
            liver disorders.  Defectives are particularly sensitive to
            pollution induced inflammation.  Injection of anti-trypsin relieves
            the condition [1].
FEATURES             Location/Qualifiers
     source          1. .1352
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="14q32.1"
     gene            1. .1352
                     /gene="PI"
     mRNA            <1. .1352
                     /gene="PI"
                     /product="a1-antitrypsin mRNA"
     CDS             20. .1276
                     /gene="PI"
                     /note="alpha-1-antitrypsin"
                     /codon_start=1
                     /protein_id="AAB59375.1"
                     /db_xref="GI:177829"
                     /db_xref="GDB:G00-120-289"
                     /translation="MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDH
                     PTFNKITPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILE
                     GLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKK
                     LYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKW
```

```
                     ERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIF
                     FLPDEGKLQHLENELTHDIITKFLENEDRRSASLHLPKLSITGTYDLKSVLGQLGITK
                     VFSNGADLSGVTEEAPLKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNK
                     PFVFLMIEQNTKSPLFMGKVVNPTQK"
     sig_peptide     20. .91
                     /gene="PI"
                     /note="alpha-1-antitrypsin signal peptide"
     mat_peptide     92. .1273
                     /gene="PI"
                     /product="alpha-1-antitrypsin mature peptide"
     old_sequence    241
                     /gene="PI"
                     /citation=[1]
     conflict        286
                     /gene="PI"
                     /citation=[3]
                     /replace=""
     old_sequence    435. .439
                     /gene="PI"
                     /citation=[1]
     old_sequence    748
                     /gene="PI"
                     /citation=[1]
     old_sequence    784
                     /gene="PI"
                     /citation=[1]
     old_sequence    787
                     /gene="PI"
                     /citation=[1]
     old_sequence    799. .802
                     /gene="PI"
                     /citation=[1]
     old_sequence    837
                     /gene="PI"
                     /citation=[1]
     old_sequence    1114
                     /gene="PI"
                     /citation=[1]
BASE COUNT      349 a    386 c    325 g    292 t
ORIGIN      90 bp upstream of DdeI site.

  Query Match           28.1%;  Score 429;  DB 9;  Length 1352;
  Best Local Similarity  60.2%;  Pred. No. 2.6e-81;
  Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || || ||    || || || || || ||
Db      92 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 151

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || ||  | || || || || ||     | ||  | ||
Db     152 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 211

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
            | || || ||    ||    || || || || ||    || ||    || || || || ||
Db     212 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 271

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || ||  | || ||  |
Db     272 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 331

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || ||  |  |
Db     332 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 391

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || ||  ||    ||  | ||  | || || || ||.||  | ||  |
Db     392 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 451

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
              || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     452 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 511

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || || |||||
Db     512 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 571

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || ||  | || || ||  | ||  | || ||
Db     572 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 631

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || ||||  | || || || || ||
Db     632 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 691

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     692 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 751

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || || || || || ||  |      |||||  |
Db     752 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 811

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || ||  | || || || || || ||
Db     812 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 871

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || ||  | || || || ||
Db     872 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 931

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           |  |    ||      | ||  | || ||  |    || || || || || ||  | ||
Db     932 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 991

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  | || ||  | || || || || || ||    || || || || |    ||
Db     992 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1051

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db    1052 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1111

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || || || || || || || || || ||
Db    1112 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1171

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db    1172 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1231

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
              ||  | || ||||| || || || || || || || ||
Db    1232 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1272


RESULT 10
AX335338
LOCUS       AX335338               1371 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 5847 from Patent WO0194629.
ACCESSION   AX335338
VERSION     AX335338.1  GI:18126057
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 5847 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1371
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       350 a    388 c    340 g    293 t
ORIGIN

  Query Match             28.1%;  Score 429;  DB 6;  Length 1371;
  Best Local Similarity   60.2%;  Pred. No. 2.6e-81;
  Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || ||    || || || || || || ||
Db     111 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 170

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || || || || || || || |  || ||  | ||
Db     171 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 230

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || || ||
Db     231 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 290

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || || |  | || ||  |
Db     291 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 350

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || ||  |  |
Db     351 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 410

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || || ||  | ||  |
Db     411 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 470

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
              || ||  | ||  | || || || ||  | || || || || ||  | || || ||
Db     471 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 530

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || || |||||
Db     531 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 590

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || ||  | || || ||  | ||  | || ||
Db     591 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 650

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || || ||||  | || || || ||
Db     651 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 710

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     711 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 770

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || || || || ||  |       |||||  |
Db     771 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 830

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
            | ||||| ||  | || || || || || || ||  | || || || || || || ||
Db     831 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 890

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || ||  | || || || || ||
Db     891 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 950

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           | |  ||     | ||  | || ||  |    || || || || || || ||  | ||
Db     951 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1010

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  | || ||  | || || || || || || ||    || || || ||  |    ||
Db    1011 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1070

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db    1071 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1130

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || |||||  | || || || || ||| ||
Db    1131 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1190

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db    1191 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1250

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
              ||  | || ||||| || || || || || || || ||
Db    1251 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1291


RESULT 11
HSATPR1
LOCUS       HSATPR1                 1371 bp    mRNA    linear   PRI 05-MAR-2002
DEFINITION  Human mRNA for alpha 1-antitrypsin.
ACCESSION   X01683
VERSION     X01683.1  GI:28965
KEYWORDS    antitrypsin; signal peptide.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 274)
  AUTHORS   Kurachi,K., Chandra,T., Degen,S.J., White,T.T., Marchioro,T.L.,
            Woo,S.L. and Davie,E.W.
  TITLE     Cloning and sequence of cDNA coding for alpha 1-antitrypsin
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 78 (11), 6826-6830 (1981)
  MEDLINE   82082539
   PUBMED   7031661
REFERENCE   2  (bases 1 to 1371)
  AUTHORS   Bollen,A., Herzog,A., Cravador,A., Herion,P., Chuchana,P., Vander
            Straten,A., Loriau,R., Jacobs,P. and van Elsen,A.
  TITLE     Cloning and expression in Escherichia coli of full-length
            complementary DNA coding for human alpha 1-antitrypsin
  JOURNAL   DNA 2 (4), 255-264 (1983)
  MEDLINE   84107980
   PUBMED   6319097
REFERENCE   3  (bases 1 to 1352)
  AUTHORS   Colau,B., Chuchana,P. and Bollen,A.
  TITLE     Revised sequence of full-length complementary DNA coding for human
            alpha 1-antitrypsin
  JOURNAL   DNA 3 (4), 327-330 (1984)
  MEDLINE   85026667
   PUBMED   6333329
REFERENCE   4
  AUTHORS   Rosenberg,S., Barr,P.J., Najarian,R.C. and Hallewell,R.A.
  TITLE     Synthesis in yeast of a functional oxidation-resistant mutant of
            human alpha-antitrypsin
  JOURNAL   Nature 312 (5989), 77-80 (1984)
  MEDLINE   85036645
   PUBMED   6387509
COMMENT     Revised sequence in Colan B., Chuchana P., Bollen A.; DNA
            3:327-330(1984).
            Data kindly reviewed (11-FEB-1986) by A. Bollen.
FEATURES             Location/Qualifiers
     source          1. .1371
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             39. .1295
                     /codon_start=1
                     /product="alpha 1-antitrypsin"
                     /protein_id="CAA25838.1"
                     /db_xref="GI:28966"
                     /db_xref="SWISS-PROT:P01009"
                     /translation="MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDH
                     PTFNKITPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILE
                     GLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKK
```

```
                     LYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKW
                     ERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIF
                     FLPDEGKLQHLENELTHDIITKFLENEDRRSASLHLPKLSITGTYDLKSVLGQLGITK
                     VFSNGADLSGVTEEAPLKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNK
                     PFVFLMIEQNTKSPLFMGKVVNPTQK"
     sig_peptide     39. .110
     mat_peptide     111. .1292
                     /product="unnamed"
     polyA_site      1371
BASE COUNT      350 a    388 c    340 g    293 t
ORIGIN

  Query Match           28.1%;  Score 429;  DB 9;  Length 1371;
  Best Local Similarity  60.2%;  Pred. No. 2.6e-81;
  Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || || ||    || || || || || ||
Db     111 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 170

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || ||  | || || || || ||     | ||  | ||
Db     171 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 230

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || ||
Db     231 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 290

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || ||  | || ||  |
Db     291 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 350

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || ||  |  |
Db     351 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 410

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || ||  | ||  |
Db     411 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 470

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
           || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     471 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 530

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || |||||
Db     531 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 590

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || || || || || || || || || || ||
Db     591 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 650

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || || ||||| | || || || ||
Db     651 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 710

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     711 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 770

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           ||||| |  | || ||||| || || || || || || || |       |||||  |
Db     771 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 830

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || || ||  | || || || || || ||
Db     831 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 890

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || ||  | || || || ||
Db     891 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 950

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
```

```
           | |   ||    | ||  | || ||  |    || || || || || ||  | ||
Db     951 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1010

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  | || ||  | || || || || || ||    || || || ||  |    ||
Db    1011 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1070

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db    1071 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1130

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || |||||  | || || || || |||   ||
Db    1131 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1190

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db    1191 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1250

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
             ||  | || ||||| || || || || || || || ||
Db    1251 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1291


RESULT 12
AK026174
LOCUS       AK026174                1399 bp    mRNA    linear   PRI 29-SEP-2000
DEFINITION  Homo sapiens cDNA: FLJ22521 fis, clone HRC12386, highly similar to
            AF113676 Homo sapiens clone FLB2803 PRO0684 mRNA.
ACCESSION   AK026174
VERSION     AK026174.1  GI:10438937
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens primary human renal epithelial cells cDNA to mRNA,
            clone_lib:HRC clone:HRC12386.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1399)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .1399
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC12386"
                     /cell_type="primary human renal epithelial cells"
                     /clone_lib="HRC"
                     /note="cloning vector pME18SFL3"
     misc_feature    1. .1399
                     /note="highly similar to AF113676 Homo sapiens clone
                     FLB2803 PRO0684 mRNA"
BASE COUNT      378 a    395 c    329 g    297 t
ORIGIN
```

```
Query Match             28.1%;  Score 429;  DB 9;  Length 1399;
Best Local Similarity   60.2%;  Pred. No. 2.6e-81;
Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || ||    || || || || || ||
Db     116 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 175

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || | || || || || || ||     | ||  | ||
Db     176 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACTGCCAG 235

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || || ||
Db     236 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 295

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || ||  | || ||  |
Db     296 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 355

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || ||  |  |
Db     356 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 415

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    || | ||  | || || || || || ||  | ||  |
Db     416 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 475

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
           || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     476 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 535

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || || ||
Db     536 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 595

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || ||  | || || ||  | ||  | || ||
Db     596 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 655

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || ||  | || || || ||
Db     656 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 715

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     716 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 775

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || || || || || ||  |       |||||  |
Db     776 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 835

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || || ||  | || || || || ||
Db     836 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 895

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || || ||  | || || || ||
Db     896 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 955

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           |  |    ||     | ||  | || ||  |    || || || || || ||  | ||
Db     956 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1015

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
           ||  | || ||  | || || || || ||    || || || ||  |    ||
Db    1016 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1075

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db    1076 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1135

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || || |||||  | || || || || |||   ||
Db    1136 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1195

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db    1196 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1255

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
              ||  | || ||||| || || || || || || || ||
Db    1256 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1296


RESULT 13
E00195
LOCUS       E00195                  1434 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  cDNA encoding human antitrypsin.
ACCESSION   E00195
VERSION     E00195.1  GI:2168491
KEYWORDS    JP 1984091886-A/1.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1434)
  AUTHORS   Guren,E.K. and Richiyaado,U.
  TITLE     SUGAR LYSING PROMOTOR : PROTEASE INHIBITOR FOR DEVELOPMENT OF
            ADJUSTED PROTEIN
  JOURNAL   Patent: JP 1984091886-A 1 26-MAY-1984;
            JIMOSU CORP
COMMENT     OS   (human)
            PN   JP 1984091886-A/1
            PD   26-MAY-1984
            PF   12-AUG-1983 JP 1983147907
            PR   13-AUG-1982 US 82    408099, 28-APR-1983 US 83    489406 PI
            GUREN EICHI KAWASAKI, RICHIYAADO UTSUDOBERII
            PC   C12N15/00,A61K35/72,A61K37/64,C07G7/00,C07H21/04,C12N1/16, PC
             C12N9/99,
            PC   C12P21/00,(C12N1/16,C12R1:865);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1. .46
            FT   CDS             47. .1303
            FT   3'UTR           1304. .1434
            FT   mRNA            1. .1434
            FT   sig_peptide     47. .118
            FT   mat_peptide     119. .1300
            FT                   /product='human antitrypsin'.
FEATURES             Location/Qualifiers
     source          1. .1434
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      369 a    432 c    340 g    293 t
ORIGIN

  Query Match             28.1%;  Score 429;  DB 6;  Length 1434;
  Best Local Similarity   60.2%;  Pred. No. 2.6e-81;
  Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || ||    || || || || || ||
Db     119 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 178

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || | || || || || || ||     | ||  | ||
Db     179 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 238

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || || ||
Db     239 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 298
```

```
Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
          || || ||| |      | || || || || || || || || || ||  | || ||  |
Db    299 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 358

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
          || || ||  | || || || || || || || || || || || || || || || ||  |  |
Db    359 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 418

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
          | ||  | || || || ||    ||  | ||  | || || || || || ||  | ||  |
Db    419 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 478

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || || ||  | || || || || || ||  | || ||
Db    479 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 538

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
          || || || || || || || || || || || || || || || || || || || |||||
Db    539 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 598

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
          || ||||| || || || || || || || || || || || || || || || || || ||
Db    599 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 658

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
          || || || || || || || || || || || || || || || || || || || ||||
Db    659 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 718

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
          || || || || || || || || || || || || || || || || || || || |||
Db    719 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 778

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
          |||||  |  | || |||| || || || || || ||  |  |||||  |
Db    779 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 838

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
          | ||||| ||  | || || || || || || || || ||  | || || || || || ||
Db    839 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 898

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
          || ||  | || |||||  | || || || || || || || ||  | || || || ||
Db    899 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 958

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
          | |   ||    | ||  | || ||   |    || || || || || || ||  | ||
Db    959 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1018

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
          ||  | || ||  | || || || || || ||    || || || || |   ||
Db   1019 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1078

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
          || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db   1079 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1138

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
          || || || || || || || || || || || || || || || || |||||  | || ||  ||
Db   1139 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1198

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
          || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db   1199 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1258

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
             ||  | || ||||| || || || || || || || ||
Db   1259 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1299

RESULT 14
I04196
LOCUS       I04196                 1434 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 3 from Patent EP 0137633.
ACCESSION   I04196
VERSION     I04196.1  GI:591838
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1434)
  AUTHORS   Parker,M.L. and Kawasaki,G.H.
  TITLE     Method of expressing alpha-1-antitrypsin in bacteria and its use in
            therapeutic formulations, and vectors and bacteria for such method
            and their production
  JOURNAL   Patent: EP 0137633-A1 3 17-APR-1985;
FEATURES             Location/Qualifiers
     source          1. .1434
                     /organism="unknown"
BASE COUNT      369 a    432 c    340 g    293 t
ORIGIN

  Query Match           28.1%;  Score 429;  DB 6;  Length 1434;
  Best Local Similarity   60.2%;  Pred. No. 2.6e-81;
  Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
          || || || || || || || || || || || || || ||     || || || || || ||
Db    119 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 178

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
          || || || || || || || || || || || || || ||     | || ||  | ||
Db    179 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 238

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
          | || || ||     ||    || || || || ||     || ||     || || || ||
Db    239 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 298

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
          || || ||| |      | || || || || || || || || || ||  | || ||  |
Db    299 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 358

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
          || || ||  | || || || || || || || || || || || || || || || ||  |  |
Db    359 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 418

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
          | ||  | || || || ||    ||  | ||  | || || || || || ||  | ||  |
Db    419 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 478

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || || ||  | || || || || || ||  | || ||
Db    479 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 538

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
          || || || || || || || || || || || || || || || || || || || |||||
Db    539 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 598

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
          || ||||| || || || || || || || || || || || || || || || || || ||
Db    599 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 658

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
          || || || || || || || || || || || || || || || ||||| | || || || ||
Db    659 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 718

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
          || || || || || || || || || || || || || || || || || || || |||
Db    719 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 778

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
          |||||  |  | || |||| || || || || || ||  |  |||||  |
Db    779 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 838

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
          | ||||| ||  | || || || || || || || || ||  | || || || || || ||
```

```
Db    839 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 898

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851

Db    899 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 958

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911

Db    959 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1018

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971

Db   1019 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1078

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031

Db   1079 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1138

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091

Db   1139 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1198

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151

Db   1199 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1258

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192

Db   1259 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1299


RESULT 15
I04272
LOCUS       I04272                  1434 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 1 from Patent EP 0139383.
ACCESSION   I04272
VERSION     I04272.1  GI:591814
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1434)
  AUTHORS   Russell,P.R.
  TITLE     Method for expressing foreign genes in schizosaccharomyces pombe
            and the use in therapeutic formulations of the products, DNA
            constructs and transformant strains of schizosaccharomyces pombe
            usable in such method and their preparation
  JOURNAL   Patent: EP 0139383-A1 1 02-MAY-1985;
FEATURES             Location/Qualifiers
     source          1. .1434
                     /organism="unknown"
BASE COUNT      369 a    432 c    340 g    293 t
ORIGIN

  Query Match           28.1%;  Score 429;  DB 6;  Length 1434;
  Best Local Similarity   60.2%;  Pred. No. 2.6e-81;
  Matches  711;  Conservative    0;  Mismatches  470;  Indels    0;  Gaps    0;

Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71

Db    119 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 178

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131

Db    179 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 238

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191

Db    239 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 298

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251

Db    299 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 358

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311

Db    359 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 418

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371

Db    419 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 478

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431

Db    479 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 538

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491

Db    539 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 598

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551

Db    599 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 658

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611

Db    659 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 718

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671

Db    719 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 778

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731

Db    779 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 838

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791

Db    839 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 898

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851

Db    899 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 958

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911

Db    959 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 1018

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971

Db   1019 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1078

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031

Db   1079 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1138

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091

Db   1139 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1198

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151

Db   1199 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1258

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192

Db   1259 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1299


Search completed: December  7, 2002, 01:14:57
Job time : 2698.5 secs
```